

A neural network approach for determining spatial and geometry dependent Green's functions for thermal stress approximation in power plant header components

J.P. Rouse^{a,*}, C.J. Hyde^a, A. Morris^b

^aDepartment of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, Nottinghamshire, NG7 2RD, UK

^bEDF Energy, West Burton Power Station, Retford, Nottinghamshire, DN22 9BL, UK

Abstract

The trend in power generation to operate plant with a greater frequency of on/partial/off load conditions creates several concerns for the long term structural integrity of many high temperature components. The Green's function method has been used for many years to estimate the thermal stresses in components such as steam headers by attempting to solve the un-coupled thermal stress problem for a unit temperature step. Once a Green's function for a unit temperature step has been determined, realistic or actual component temperature profiles can be discretised and the time dependent stress profile reconstructed using Duhamel's theorem. Stress fluctuations can therefore be estimated and damage due to fatigue mechanisms can be quantified. A potential difficulty with this method is that Green's function approximations are determined for a single analysis point in a structure. This is because Green's functions are approximated by fitting a trial function to the results of finite element (FE) simulations. While a user can make some judgement on which point in a structure will give the "worst case" (or life limiting) conditions, it is foreseeable that points of interest will be dependent on the specific analysis conditions, such as the stub penetration geometry and the loading condition considered. The neural network approach described in this paper provides a means where transient thermal stress models of complex components (here taken to be steam headers) can be generated relatively quickly and used pro-actively to assess and modify plant operation. A range of header geometries have been considered to make the network applicable over an industry relevant envelope. Coefficients of determination (R^2) are typically above 0.92 when reconstructed (from neural network results) unit temperature step stress profiles are compared against "true" FEA results. Mean errors in the stress profiles are, for the majority of cases, less than

[☆]This document is a collaborative effort.

*Corresponding author

10%. Suggestions are also made on possible future improvements to the method through the use of additional constraints on the reconstructed stress profiles.

Keywords: Power plant, Header, Green's Function, Thermal Stress, Neural Network

1. Introduction

Pressure is mounting on power plant operators to generate electricity in an efficient and economical manner. Unit loads will fluctuate with higher frequencies and steeper "ramp up and down" rates as drivers attempt to match market demands. As steam pressures and temperatures vary with time, potentially large thermal stresses will develop in thick walled components such as steam headers. These thermal stresses will be necessarily time dependent and, should constant operation be maintained for a sufficient period of time, will tend to zero as the component reaches a homogeneous temperature. The fluctuation of total stress in components makes fatigue an important structural integrity concern in power plant components; a problem which is significantly complicated by the transient nature of thermal stresses. Large conventional power stations were originally designed for limited thermal cycling, however in response to the commercial environment it is not uncommon now to find these large stations operated with up to 250 Unit starts per year, which results in the regular occurrence of fatigue cracking and creep-fatigue cracking.

Many novel monitoring systems have been developed for assessing the structural integrity of at risk power plant components, including "on line" management systems that monitor power station load characteristics (such as main steam temperature and pressure) and estimate component degradation using generalised finite element models and creep/fatigue damage fraction rules[1, 2, 3, 4]. An example of one of these products is Areva's fatigue monitoring system FAMOSi[5, 6], where thermal loads are recorded using on site thermocouples and converted to thermal stresses using FEA models at critical points in a system. Alternatively, accurate stress histories in a component may be estimated through bespoke analyses utilising complex visco-plastic material models[7], however this is commonly computationally intensive and is typically impractical for on line component assessment. In the UK, the approach to managing the integrity of boiler steam headers is based primarily on an Inspection Based Assessment (IBA). This approach is implemented using a range of non-destructive inspection based assessment techniques during the four-yearly statutory outages. The extent/scope of site outage inspections for boiler steam headers is prescribed in various industry guidelines, and threats to the integrity of boiler steam headers can arise from, erosion damage, creep rupture, thermal fatigue, creep fatigue and corrosion fatigue.

While these advances have shown some success, established design codes and analysis procedures are still by far the most commonly used tools in industry for component fitness assessment, along with frequent inspection during

outage periods[8]. In the UK, the R5[9, 10] procedure is commonly used for high temperature assessment and the R6[11] procedure for low temperature fracture assessment of power plant components. These step by step methods usually involve decomposing a loading history into cycles. The likelihood of failure by various mechanisms, such as plastic collapse, creep and fatigue, is calculated by estimating damage accumulation and mechanism interaction factors.

The Green's function method provides a general approach to estimate the transient linear elastic thermal stress responses at a point in a structure by integrating the response due to a unit thermal load change. In the context of steam headers, thermal stress histories may be estimated at a point of interest for any bulk steam temperature history. While limited to linear analysis, the Green's function method is still of use in component failure assessment, particularly where damage is suspected to be localised. The Green's function method (see section 2) has been shown to be a useful tool in predicting transient thermal stresses by several authors. In particular, the technique has been applied to fatigue analysis problems in the nuclear power industry[12, 13, 14].

Boiler header systems transfer and distribute water/steam between various boiler stages, hence operate at temperatures both below and above the creep limit for typical conventional plant materials, which is approximately 400°C. Maximum design conditions for these header systems on large conventional power stations are typically 568°C and 170 bar. Typical areas for inspection on a boiler header during an outage includes the following:

- End cap welds regions
- Stub-to-Shell fillet welds and stub pipes
- Branch connection welds
- Header body internals (by remote visual inspection)
- Shell construction butt welds (in larger header components)
- Any attachment welds on the header shell (support connections, for example)
- Header shell parent material and inter-stub pipe ligament positions
- Header shell diametral measurements (for creep strain monitoring)

Operational steam temperatures are usually used to provide an initial assessment of the rate of creep life consumption, which is often supplemented later in life by the installation of surface mounted thermocouples on the header shell and selected stub tube locations. Clearly as the components age the extent of invasive assessments during a statutory outage increases and a significant amount of metallurgical data on header condition can be accumulated. This metallurgical data is not specifically used in a forward life prediction however; it is usually used to provide confidence that the next 4-year operating period can be safely navigated before repeat inspections. It is clearly advantageous if

more predictive models of life consumption can be used, moreover, as a means of directing the plant operators to adjust operation and hence reduce the rate of header life consumption where needed. The neural network approach described in this paper provides a means where transient thermal stress models of these complex components can be generated relatively quickly and used pro-actively to assess and modify plant operation.

2. Background to the Thermoelastic Problem and Neural Networks

2.1. The Linear Thermoelastic Problem and Coupling

The governing equations for a linear coupled thermoelastic problem may be derived from the fundamental principles of mechanics and thermodynamics. When loads applied to a body give rise to variations in strain within the body, variations in temperature are also observed. This causes heat flow and therefore an increase in entropy for the body (this irrecoverable mechanical dissipation is known as thermoelastic dissipation). There is an internal generation of heat due to mechanical deformation that will affect the temperature field within a body in addition to any thermal boundary conditions. Deformation however is not only controlled by the application of, say, body forces. Temperature fields cause thermal expansion within elements of the body, generating additional internal surface forces between the elements. There exists therefore a coupling between the solutions for temperature and displacement fields, $T(P, t)$ and $\mathbf{u}(P, t)$ respectively (where P is a point within the body specified by coordinates using the coordinate system $x_1, x_2, x_3 = \mathbf{x}$ and t is time). For a linear coupled thermoelastic problem, it can be shown that a unique solution may be found (for a given set of initial and boundary conditions) using the heat equation with mechanical coupling (equation (1)), the equilibrium condition (equation (2)), noting the inclusion of an inertia term), the strain-displacement relations (equation (3)), and the stress-strain relations (equation (4))[15]. Note the use of indicial notation ($\frac{\delta g_i}{\delta x_j} = g_{i,j}$, where g_i is a vector component in the i^{th} direction and x_j is the basis vector in the j^{th} direction of the coordinate system, $i, j = 1, 2, 3$) and the Einstein summation convention. Note also that dots are used to denote derivatives with respect to time.

$$kT_{,mm} = \rho c \dot{T} + (3\lambda + 2\mu) \alpha T_0 \dot{\epsilon}_{kk} \quad (1)$$

$$\sigma_{ij,j} + f_i = \rho \ddot{u}_i \quad (2)$$

$$\epsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \quad (3)$$

$$\sigma_{ij} = \delta_{ij} \lambda \epsilon_{kk} + 2\mu \epsilon_{ij} - \delta_{ij} (3\lambda + 2\mu) \alpha T \quad (4)$$

where ϵ , σ , and f are the small strain tensor, stress tensor and body force vector field, respectively. T_0 is a reference temperature at which, in the absence of body forces, the material will be in a stress free state. Material dependent parameters are thermal conductivity (k), density (ρ), specific heat capacity at constant deformation (c), and the thermal expansion coefficient (α). Lamé's first and second parameter are defined in terms of Young's modulus (E) and Poisson's ratio (ν) in equation (5). It has been assumed in the present work that any temperature dependency in these material parameters is negligible. δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, else $\delta_{ij} = 0$).

$$\begin{aligned}\lambda &= \frac{Ev}{(1+\nu)(1-2\nu)} \\ \mu &= \frac{E}{2(1+\nu)}\end{aligned}\tag{5}$$

The existence of the coupling term in the energy equation (equation (1)) greatly complicates the solution process for the thermoelastic problem (clearly temperature and displacement field solutions must be found simultaneously to satisfy equations (1) to (4) and the problem specific initial and boundary conditions). In general, temperature variations due to mechanical deformations are small (particularly if small strain theory is being implemented). Similarly, differences between heat transfer solutions in deformed and undeformed bodies are also small (deformations from either thermal expansion or external mechanical agencies do not change the dimensions of the structure to such an extent that heat transfer is significantly affected). If thermoelastic dissipation can be neglected (conditions for this are given in section 3.3), an uncoupled formulation may be derived for the thermoelastic problem. In this case, internal heat generation due to deformation is ignored and temperature fields can be found first by solving the well know heat equation (equation (6)), where κ is the thermal diffusivity ($\kappa = k/\rho C_p$). Once the temperature field has been determined, the corresponding displacement field (dependent on thermal expansion and mechanical deformation) may be found.

If the rate of change of the deformation rates are small (as is the case in many engineering applications) inertia effects may be neglected and the formulation is termed quasi-static. In this case (with the absence of body forces) the equilibrium equation simplifies to Equation (7). Strain-displacement and stress-strain relations given in equations (3) and (4), respectively, are still valid in the uncoupled formulation.

$$\kappa T_{,mm} = \dot{T}\tag{6}$$

$$\sigma_{ij,j} = 0\tag{7}$$

2.2. The Green's Function Method for Estimating Transient Thermal Stresses

The thermoelastic problem has been defined in section 2.1. Solutions for even the uncoupled formulation with the component geometry used in the

present work (power plant steam header) are very complex and generally require numerical methods (such as finite element analysis, FEA) to estimate a solution. A practical problem with this analysis strategy is that temperature, displacement, and consequently stress fields are required for each new operating condition. It is not feasible to perform full FEA simulations of the header components for each change to the bulk steam temperature or pressure. A solution to this dilemma however exists through the use of Green's functions.

It can be seen from section 2.1 that the thermal stress solution is based on the temperature field solution, both of which are unique and dependent upon the particular initial and boundary conditions of the problem. This uniqueness allows the use of a Green's function that finds the thermal stresses based on the boundary conditions. It is therefore possible to determine the thermal stress distribution without direct knowledge of the temperature or displacement fields. As the present work is concerned with a power plant header application, the bulk internal steam temperature may be used as a "driving" term for the thermal stress field (it shall be assumed that external surfaces of the header are insulated and attention is restricted to the uncoupled formulation). Thermal stresses at a point P in the header structure can be found by the integral in equation (8), where $G(P, t - \tau)$ is the Green's function, $\psi(t)$ is the bulk internal steam temperature, and τ is the time integration variable.

$$\sigma(P, t) = \int_0^t G(P, t - \tau) \frac{d\psi(\tau)}{d\tau} d\tau \quad (8)$$

Numerical integration of equation (8) can be achieved using equation (9), where $G_{SS}(P) = \lim_{t \rightarrow \infty} G(P, t)$. Note that in the present work, with a steady state constant bulk steam temperature and in the absence of mechanical body forces, the asymptotic value of G is 0. A characteristic time (t_{CH}) is used in equation (9) to specify a limit for the time integral. It is defined as the point at which the stress contribution predicted by $G(P, t)$ decays to a value which is considered to be negligible. Here, t_{CH} is taken as the time when $G(P, t)$ is 5% [14] is the maximum value ($\hat{\sigma}$). An exponential sum of terms (see equation (10)) is used in the present work to estimate unit temperature step stress responses.

$$\sigma(P, t) = G_{SS}(P)\psi(t) + \sum_{t-t_{CH}}^t \bar{G}(P, t - \tau) \Delta\psi(\tau) \quad (9)$$

$$\bar{G}(P, t) = \exp\left(\sum_{m=1}^7 C_m(P) (\ln(t))^{m-1}\right) \quad (10)$$

2.3. An Overview of Neural Networks

Neural networks were first proposed by McCulloch and Pitts in 1943 [16] and have been implemented in literature to determine patterns and trends in data where some correlation exists but controlling functions are either unknown or too complex to implement [17, 18]. Multiple authors have used neural networks

for structural integrity assessments of specific materials and/or components. For example, Venkatesh and Rack used a back propagation neural network for the prediction of service life of the Nickel based alloy INCONEL 690 operating under creep fatigue interaction conditions[19]. Structural integrity assessments have also been approximated by trained neural networks for specific component types, including fatigue damage in trabecular bone structures[20] and contact fatigue in alloy cast steel rollers [21]. More recently, Kao and Yeh have used neural networks to improve the integration between structural solvers (FEA packages, for example) and optimisation routines[22] by determining characteristic responses for simple two truss element structures. Stress analysis in steel catenary risers was studied using neural networks in the work of Quéau et al. in order to inform fatigue calculations[23].

Artificial neural networks are mathematical representations of biological central nervous systems[24]. A brief overview of neural networks will be given here, with detailed descriptions available from numerous sources[24, 25]. Fundamentally, a neural network neuron takes a group of inputs (or signals in a vector X), multiplies them by weighting values (W_i) and pass them through mathematical functions ($S(X_i, W_i, \theta)$), giving rise to the effective input S . This function is known as the propagation rule and, while many variants exist, the most simple and applicable is the weighted summation rule in equation (11) (shown for a single neuron fed by i inputs). External inputs (or biases, b) may be used to distinguish certain neurons in a network. The effective input is used in an activation function $F(S)$ to determine the neuron's output Y .

Weights and biases used in a particular network are determined in a process called training. In effect, this is a form of optimisation and similar methods (such as a least squares evaluation of the Gauss-Newton method[26]) may be used for this purpose. Sets of inputs with corresponding known ("true") outputs are collected and fed into the neural network. From a usually randomised initial state, the training algorithm will iteratively alter the network's weights and biases to minimise the difference between approximated outputs from the network (Y) and the corresponding true values. Typically, a validation sequence would also be required for the approval of a network. In some cases, neural networks can "overfit" the training data set. This means that while errors in the training set are small, predicted values for inputs outside the training set (where clearly a predictive neural network has the greatest value) show very large errors.

$$S = \sum_{i=1}^n X_i W_i + \theta \quad (11)$$

For all but the simplest systems, fully connected neural networks will comprise of multiple layers of neurons that are linked together (see figure 1). Back propagated neural networks are simple examples of these feed forward networks. An array of inputs X (length m) will form the input layer. These inputs are fed to the first "hidden" layer of neurons (numbering n in figure 1). Each neuron is fed by all of the inputs (weighted in some way, depending on the

prorogation rule used) and the effective input is determined by the neuron's function S_{1j} ($1 \leq j \leq n$). An output for that neuron is determined from the effective input using the activation function. The outputs of the neurons from this first hidden layer are fed as inputs into the neurons of the next hidden layer. Commonly, local outputs from the last hidden layer of inputs are summed in an output neuron, giving rise to the network's output \hat{Y} [25]. Y is therefore the local output from a neuron and \hat{Y} is the global output from the neural network.

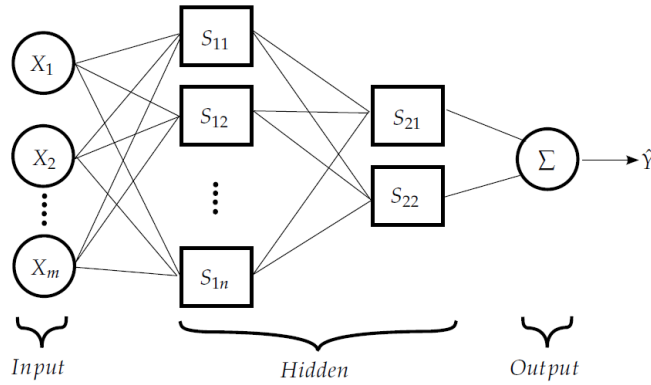


Figure 1: A 2 layer feed forward fully connected neural network[24].

Training is completed for a neural network in order to determine the magnitude of the weights and biases for each neuron. This is done for a sample training set of data. In back propagation networks, data flows uni-directionally from the input nodes to the output. During training, errors (E) between the outputs predicted by a neural network (with a particular set of weight and bias values) and the true training values are compared using the sum of squares approach (or similar). Using a gradient descent optimisation method, the change to a weight is determined by a multiplier term and the rate of change in E with respect to the weight W_i (see equation (12)). If training is undertaken for a specific data set only, it is termed supervised learning. Un-supervised learning results in weights and biases being continuously updated as new information is made available. While supervised learning is considered more relevant to the present work as the intended application demands a deterministic capability, unsupervised learning may be applicable in the future to update networks as new data (either from additional FEA or plant inspections) becomes available.

$$\Delta W_i = -\eta \frac{\delta E}{\delta W_i} \quad (12)$$

The example neural network shown above is an example of a fully connected network. This is not an accurate representation of a biological neural network and tends to lead to significant redundancy in the network (i.e. storage, computation and training requirements are far greater for fully connected

networks[27]). Customised networks can be developed that connect only specified neurons, meaning that the network can (potentially) be more efficient. While only feed forward neural networks are described here, recurrent topologies have been used in literature[24], such as the Boltzmann and Hopfield networks. In these cases, information does flow in one direction only but can cycle on a local basis, potentially allowing for more complex functions to be analysed. These aspects are considered outside the scope of the present work but could be implemented in the future for more sophisticated analysis networks.

Initial values for the weights and biases used in the network must be found before the training procedure begins in order to provide a “starting point”. Initialisation is a critical step in the development of any neural network and determining appropriate methods has been the subject of several investigations by many authors[28, 29, 30]. Improper initialisation can lead to numerical difficulties which can dramatically affect the performance of a neural network. For example, premature saturation occurs when activation functions output saturated values (0 or 1) in the early stages of training. This has the effect of trapping the network at a non-optimised state, potentially severely increasing training time and reducing the likelihood convergence[31]. Alternatively, solutions may converge on a local minimum[32].

In the present work, randomised initial conditions were assumed for all neuron layers using the Nguyen-Widrow algorithm. Using this method, initial values for weights and biases take a random value (generated between the limits -1 and 1). The factors β and $norm$ (equations (13) and (14), respectively) are then calculated and used to determine a revised weight and bias (based on the randomised value, see equation (15))[33].

Note that in equations (13), (14) and (15), h is the number of hidden neurons for a particular layer, I is the total number of inputs to that layer, w_i is the i^{th} weight (or bias as the two are initiated simultaneously) determined by a random number generator and w'_i is the weight (or bias) w_i that has been adjusted by the Nguyen-Widrow algorithm. If the values of the weights and biases were plotted on a histogram for a suitably large network, a distribution would be observed with higher frequencies of weights and biases at certain values and very low frequencies at the limits of the range. During training, the high frequency regions may shift and redistribute, however by localising them to begin with a faster training rate is observed than if a “hard” (uniform) or Gaussian distribution had been implemented. For the remainder of the neuron layers, zero initial values for the weights and biases were assumed (meaning that these neurons are not active, at least initially). During the training procedure, non-zero values are quickly determined in these neurons. The initialisation procedure described is not dissimilar to that used by Yoon et. al [34], whereby weights on some input nodes were artificially increased before training to emphasise those inputs the authors suspected would particularly influence a results.

$$\beta = 0.7h^{1/I} \tag{13}$$

$$norm = \sqrt{\sum_{i=0}^{i=I} w_i^2} \quad (14)$$

$$w'_i = \frac{\beta w_i}{norm} \quad (15)$$

3. Methodology overview

3.1. Unit response critical parameters and profile reconstruction

As would be expected, general features are observed for all thermal stress responses due to unit temperature steps. The precise shape of the thermal stress profile at a given point is dependent upon a structure's geometry and material, however stresses will typically quickly increase (non-linearly) to a maximum value ($\hat{\sigma}$) before decaying to an asymptotic value (zero if the boundary conditions are such that a homogeneous temperature field is permitted and in the absence of body forces). With a view to generalising the thermal stress profile so that the effect of a structure's geometry can be accounted for, three so-called "critical parameters" will now be defined. These critical parameters are values determined from the thermal stress profile and characterise its shape (see figure 2). Knowledge of these parameters will allow for the reconstruction of an approximate unit thermal stress profile (see figure 3). The critical parameters are defined as the maximum stress value ($\hat{\sigma}$, with units of *MPa* in the present work), the time at which the maximum stress is achieved ($t_{\hat{\sigma}}$, units of *s* here), and a characteristic time (t_{CH} , taken to be the time at which the instantaneous stress value is 5% of $\hat{\sigma}$ and having units of *s* here). It has been noted that defining an asymptotic value for the Green's function approximation can reduce the computational effort required to implement the method[14]. It is these critical parameters that will be used as outputs for the developed neural network. In order to achieve this, true values for the critical parameters are required from FEA simulations (discussed in section 3.3). Values for the critical parameters can be easily determined from FEA profiles using a spline approximation.

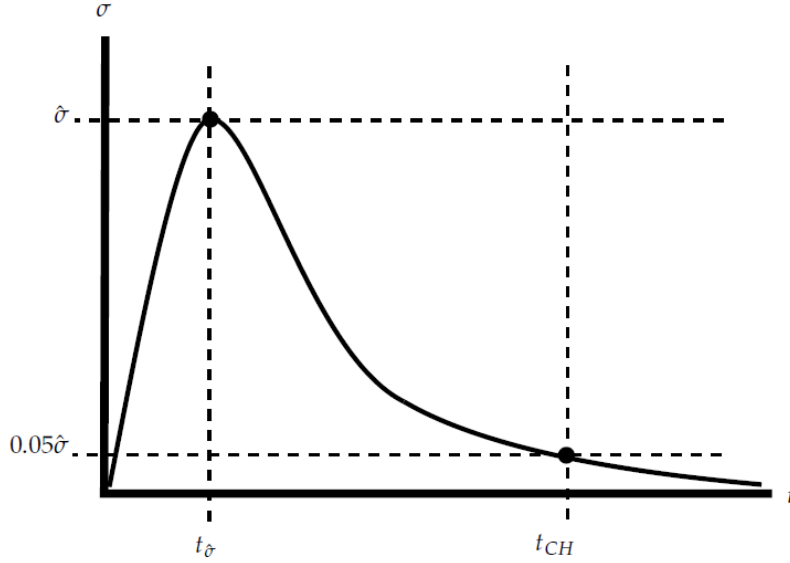


Figure 2: Critical parameters definition on the thermal stress response due to a unit temperature step, namely maximum stress $\hat{\sigma}$ (MPa), time to maximum stress $t_{\hat{\sigma}}$ (s), and characteristic time t_{CH} (s).

Once the values of the critical parameters have been determined from a trained neural network, the unit thermal stress response and Green's function approximation must be determined so that stress histories due to arbitrary bulk steam temperature profiles can be determined. In the present work, this is achieved by creating a set of intermediate points to approximate the unit stress response based on the critical parameters (see figure 3). Quadratic curves are fitted to the "ramp up" and "ramp down" regions to generate these intermediate points, with a linear approximation used in the asymptotic region. For clarity, the quadratic approximations are summarised in equations (16) and (17), where A , B and C are used as place holders for fitting coefficients (clearly it is trivial to determine these once the critical parameters are known). After intermediate points have been generated the Green's function approximation (in the present work taken as equation (10)) can be fitted (here using a Levenberg-Marquardt algorithm through LSQNONLIN in MATLAB[35]); acting as a smoothing function to remove discontinuities in the piecewise approximation and returning the reconstructed stress profile.

Three points are required to define the quadratic functions used to generate intermediate points. Two limits for each function can be found directly from the critical parameters themselves, however the third (midpoint) must be estimated. Stresses at the times $t_{\hat{\sigma}}/2$ and $t_{\hat{\sigma}} + (t_{\hat{\sigma}} - t_{CH})/2$ would be $t_{\hat{\sigma}}/2$ and $0.475t_{\hat{\sigma}}$, respectively, if a linear relationship was assumed. These values are multiplied by the factors f_1 and f_2 , respectively. It can be shown that the limits of the quadratic curve factors are $1 \leq f_1 < 2$ and $1 \geq f_2 > (2/19)$. By analysing

coefficients of determination between a random set of true (FEA) thermal stress responses and corresponding neural network derived reconstructions, values of $f_1 = 1.9$ and $f_2 = 0.48$ were determined for the reconstruction curve factors.

$$\begin{bmatrix} \frac{t_{\hat{\sigma}}}{4} & \frac{1}{2} \\ t_{\hat{\sigma}} & 1 \end{bmatrix} t_{\hat{\sigma}} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} f_1 \\ 1 \end{bmatrix} \hat{\sigma} \quad (16)$$

$$\begin{bmatrix} \hat{\sigma}^2 & \hat{\sigma} & 1 \\ \left(t_{\hat{\sigma}} + \frac{t_{\hat{\sigma}} - t_{CH}}{2}\right)^2 & t_{\hat{\sigma}} + \frac{t_{\hat{\sigma}} - t_{CH}}{2} & 1 \\ t_{CH}^2 & t_{CH} & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 \\ 0.475f_2 \\ 0.05 \end{bmatrix} \hat{\sigma} \quad (17)$$

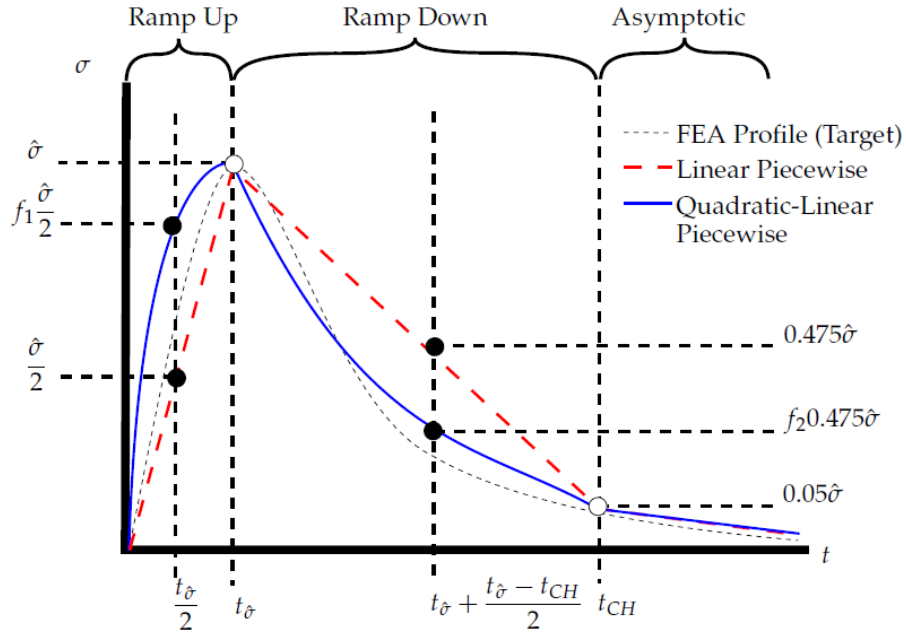


Figure 3: The stress reconstruction method, showing the definition of the curve factors f_1 and f_2 .

3.2. Neural network estimation of critical parameters

The neural network architecture used in the present work is summarised in figure 4. In essence, three networks are used in parallel to approximate the critical parameter outputs $\hat{\sigma}$, $t_{\hat{\sigma}}$, and t_{CH} . It can be proven, using the universal approximation theorem, that only one layer of hidden neurons is sufficient to approximate any function with a finite number of discontinuities (providing activation functions of the hidden units are non-linear)[25]. To this end, only one layer neural networks have been implemented for each critical parameter

in the present work through the MATLAB neural network toolbox[36]. One additional layer of hidden neurons is implemented for the calculation of $t_{\hat{\sigma}}$ in order to combine input sources as both geometrical parameters and other critical parameter ($\hat{\sigma}$ and t_{CH}) approximations are used. Inputs for the neural networks are largely derived from the geometry of the header and the spatial coordinates of a point of interest. Header geometry is defined by the diameter and thickness of the shell and stub sections (ϕ_{SHELL} , t_{SHELL} , ϕ_{STUB} and t_{STUB} , respectively). Assuming the Cartesian coordinate system shown in figure 6 with its origin on the axis of the shell section, the Cartesian spatial coordinates (x , y and z) of points around the stub penetrations may be transformed into fractions through the stub wall thickness based on a cylindrical coordinate system. These fractions fully define the location of a point of interest for a particular header geometry and are designated r_f , θ_f and z_f (see equation (18), equation (19) and equation (20)). It is generally advisable to normalise inputs prior to processing in neural networks to ensure accurate representation of the true behaviour[25], therefore all inputs have been normalised against the maximum values observed in the tested range (note r_f and z_f are normalised through their definition in equations (18) and (20)).

$$r_f = \frac{2\sqrt{x^2 + z^2}}{\phi_{STUB}} \quad (18)$$

$$\theta_f = \left| \tan^{-1} \frac{z}{x} \right| \quad (19)$$

$$z_f = \frac{2y}{\phi_{SHELL}} \quad (20)$$

It has been observed in the present work that, for a particular header geometry, correlations exist between $t_{\hat{\sigma}}$, $\hat{\sigma}$ and t_{CH} (example plots that demonstrate this are given in section 4.3). This is exploited in the present work by determining two critical parameters ($\hat{\sigma}$ and t_{CH}) using only header geometry data first and then using these with the geometry data to predict the final critical parameter ($t_{\hat{\sigma}}$). 30 hidden neurons with a hyperbolic tangent sigmoid transfer function (see equation (21) where n is the input of the neuron) and logarithmic sigmoid transfer function output neurons (see equation (22)) were found to be sufficient for the prediction of $\hat{\sigma}$ and t_{CH} . $\hat{\sigma}$ and t_{CH} are then used as inputs (along with the header geometry data) in a separate network to predict $t_{\hat{\sigma}}$. 30 hidden hyperbolic tangent sigmoid transfer function neurons are used in this network with a single logarithmic sigmoid transfer function output neuron. Training was achieved through back-propagation using a Levenberg-Marquardt algorithm and sum of square error metric[36]. Logarithmic sigmoid transfer functions were used for all output neurons as, by definition, all critical parameters presented here are positive. This will always be the case for $t_{\hat{\sigma}}$ and t_{CH} , however it is clear that the choice of stress component considered for $\hat{\sigma}$ may necessitate negative values. Equivalent von Mises stresses are the focus of the present work therefore this concern is not relevant here, however it should

be revisited if a more general analysis method is sought. Although strategies and optimisation methods for the construction of neural networks have been proposed (such as the procedure detailed by Weymaere and Martens[28] which utilises a quantification of network complexity), neuron distributions in the present network were determined using a trial and error approach. Initial values for the number of hidden neurons were determined by initialising a network with a set number of neurons, training for a limited number of iterations (100), and observing trends in mean squared error values in each of the three critical parameter predictions. Mean squared error (*MSE*) may be expressed by equation (23), where x_i is one of n reference/true observations and \hat{x}_i is the corresponding prediction. It is clear that training for a fixed number of iteration will not yield an optimum set of network weights and biases, however this was not the aim of the initial investigation. The motivation here was to gain some understanding of how the number of neurons affects the predictive capability of the network without undertaking lengthy studies. *MSE* values were determined for the training data set only. The network architecture is such that t_{δ} evaluations depend upon predicted values for the other two critical parameters. Consequently, when neuron numbers were varied for the hidden layer associated with t_{δ} , neuron numbers in the δ and t_{CH} hidden layers were fixed at 20. It should of course be noted that this method is not deterministic as initial values for the neuron weights and biases are random (as discussed previously). While the results (given in figure 5) are not strictly monotonically decreasing, a general trend may be observed and a plateau in the data emerges for all critical parameters at 20-30 neurons. As such, hidden layers were initial given 20 neurons each, with final values (given in figure 4) determined by observing when consistent total network performance across all critical parameters was achieved.

$$\frac{2}{1 + e^{-2n}} - 1 \quad (21)$$

$$\frac{1}{1 + e^{-n}} \quad (22)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (23)$$

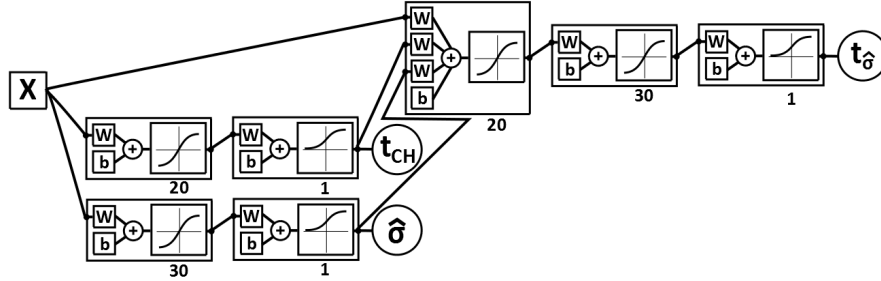


Figure 4: The neural network architecture applied in the present work, showing neuron quantities and transfer function assignments. Note input parameters (ϕ_{SHELL} , t_{SHELL} , ϕ_{STUB} , t_{STUB} , r_f , θ_f , and z_f) are summarised by the vector X . Outputs are designated by circled symbols.

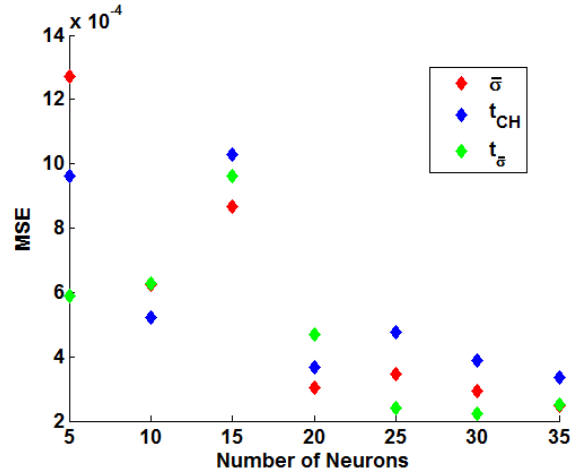


Figure 5: A summary of the (normalised) mean square errors (MSE) determined in initial neuron number investigation. Results for all critical parameters ($\hat{\sigma}$, $t_{\hat{\sigma}}$, and t_{CH}) are presented.

3.3. FEA models

FEA models must be generated in order to determine the coefficients in equation (10) and thus define the Green's functions. FEA has been conducted in the present work using the commercially available code ABAQUS. The effects of multiple stub penetrations are not considered here (although future work may look to include these effects through additional inputs). A simplified header geometry is therefore assumed that consists of a single stub penetration (see figure 6, noting the plane of symmetry assumed through the stub

penetration). Dimensions of the shell (diameter and wall thickness, ϕ_{SHELL} and t_{SHELL} , respectively) and the stub (diameter and wall thickness, ϕ_{STUB} and t_{STUB} , respectively) are varied in the present work to determine the effects of header geometry on spatially dependent Green's functions. A summary of the dimensions used in the present work (with labels) are given in tables 3 and 4.

Despite the simplified geometry, shell and stub dimensions are similar to those found in industry for P91 header components. Uncoupled thermoelastic analysis was conducted by first determining a temperature field from a heat transfer simulation. An insulated exterior boundary condition was assumed ($\dot{q} = 0$) to allow temperature fields in the model to reach equilibrium after the bulk steam temperature experiences a step change. Heat conduction on the inside surface of the header is controlled by convection (see figure 6 (b)), where the heat transfer coefficient h is taken to be a temperature independent constant $0.002W/mm^2K$ (this value is similar to that used in the work of Farraghera et al.[7]). Once the transient temperature field has been determined it can be used as an input in mechanical analyses to estimate thermal stress histories. Boundary conditions for the mechanical analysis can be seen in figure 6 (c). Tetrahedral quadratic elements were used, namely DC3D10 for thermal analyses and C3D10 for mechanical analyses (see figure 6 (a) for an example mesh)[37].

A single material is assumed for the FEA model in the present work (variations in material properties at the stub weld are not considered). Temperature dependent material parameters are required in order to calculate transient thermal stresses within the header models. Values for Young's modulus (E) and the thermal expansion coefficient (α) have been determined from monotonic tests performed on an Instron 8862 thermomechanical fatigue machine (operating under isothermal conditions) utilising radio frequency induction heating and using a TA instruments Q400 thermomechanical analyser, respectively (see table 1). Tested temperature ranges were chosen to represent the typical bounds of operation for thermal power plant components. The remainder of material constants have been taken from the work of Yaghi et al.[38] (table 2). A negligible dependency is assumed in density (ρ) and Poisson's ratio (ν) over the tested temperature range. As such values for these quantities are taken to be $7.76 \times 10^{-6} kg/mm^3$ and 0.3, respectively[38].

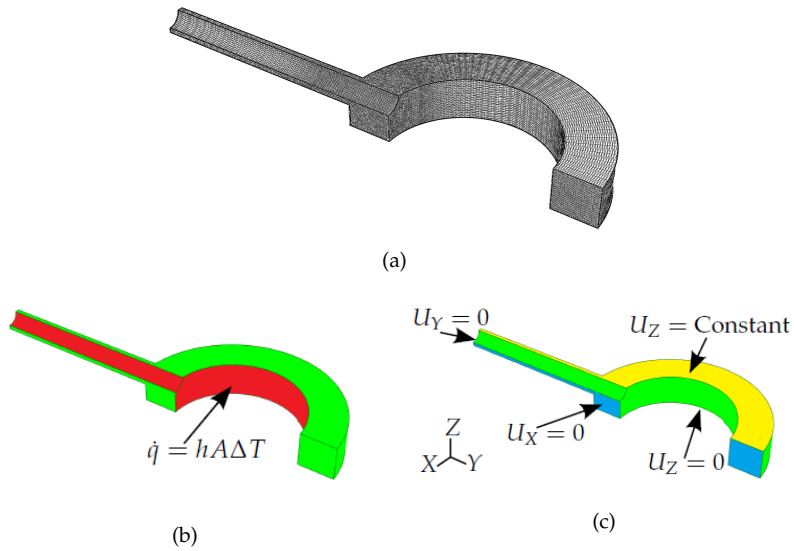


Figure 6: Finite element analysis (FEA) models, showing (a) the tetrahedral mesh, exploiting the plane of symmetry through the stub penetration, (b) boundary conditions in the thermal analyses and (c) boundary conditions in the mechanical analyses.

Table 1: A summary of the temperature dependent material parameters (representative of a P91 chrome steel), determined through experimental analysis, used in the FEA modelling.

Temperature (°C)	Young's Modulus (E) (N/mm ²)	Thermal Expansion Coefficient (α) (1/K)
240	1.88×10^5	1.21×10^{-5}
400	1.77×10^5	1.34×10^{-5}
550	1.51×10^5	1.41×10^{-5}
600	1.30×10^5	1.42×10^{-5}
700	8.58×10^4	1.41×10^{-5}

Table 2: A summary of the temperature dependent material constants (representative of a P91 chrome steel), taken from the work of Yaghi et al.[38], used in the FEA modelling.

Temperature (°C)	Thermal Conductivity (W/mm.K)	Specific Heat Capacity (C) (J/kg.K)
200	0.028	510
250	0.028	530
300	0.028	550
350	0.029	570
375	0.029	585
400	0.029	600
450	0.029	630
500	0.03	660
550	0.03	710
600	0.03	770
650	0.03	860
700	0.0305	942

Table 3: A summary of the header dimensions used to generate FEA models for training and validation of the neural network.

Header Model Label	ϕ_{SHELL} (mm)	t_{SHELL} (mm)	ϕ_{STUB} (mm)	t_{STUB} (mm)
A1	322.22	56.15	50	7
C1	435.19	41.54	50	7
E1	357.41	31.15	50	7
F1	183.33	31.15	50	7
G1	522.22	25	50	7
H1	357.41	25	50	7
D1	300	41.54	50	7
D2	300	41.54	38	6
D3	300	41.54	50	8
D4	300	41.54	70	16
D5	300	41.54	51	5
D6	300	41.54	76	8
B1	522.22	45.38	50	7
B2	522.22	45.38	38	6
B3	522.22	45.38	50	8
B4	522.22	45.38	70	16
B5	522.22	45.38	51	5
B6	522.22	45.38	76	8
I1	183.33	25	50	7
I2	183.33	25	38	6
I3	183.33	25	50	8
I4	183.33	25	70	16
I5	183.33	25	51	5
I6	183.33	25	76	8
A7	322.22	56.15	65	12
B7	522.22	45.38	65	12
C7	435.19	41.54	65	12
D7	300	41.54	65	12
E7	357.41	31.15	65	12
F7	183.33	31.15	65	12
G7	522.22	25	65	12
H7	357.41	25	65	12
I7	183.33	25	65	12

Table 4: A summary of the header dimensions used to generate FEA models for example application/extrapolation of the neural network.

Header Model Label	ϕ_{SHELL} (mm)	t_{SHELL} (mm)	ϕ_{STUB} (mm)	t_{STUB} (mm)
T1	436.99	47.08	42.89	6.44
T2	210.35	37.66	73.78	14.19
T3	282.47	54.44	39.48	9.91
T4	324.5	45.36	64.81	15.14
T5	420.48	28.22	67.1	13.47
T6	260.97	49.32	64.64	7.37
T7	454.1	46.73	53	6.55
T8	441.16	28.36	71.95	6.24
T9	298.81	31.43	55.85	7.47
T10	261.62	54.08	61.39	7.44

4. Results

4.1. Header models

Crack initiation and growth at the inner bore of a header penetration or at the ligament region between penetrations is a concern when operating thick walled components and has been observed in industry[39]. The behaviour of the region around the stub penetration is therefore of far more interest (both industrially and analytically) due to the localised thermal stress concentrations. Behaviour remote from the stub penetration is largely uniform and uncoupled thermoelastic analysis can be conducted analytically. With these considerations in mind, the neural network presented here looks to model the region of the header in the vicinity of the stub penetration (see figure 7).

For each header geometry, a random selection of 350 spatial coordinates were selected from the sampling region (see figure 8). Readers should note that figure 8 is presented here to illustrate general trends in the critical parameters between simulation cases. Transient thermal stress profiles were extracted from FEA output databases and critical parameters determined (see figure 8). It is these results that are used to train and validate the developed neural network. Spatial distributions of the critical parameters (for all header geometries given in table 3) are illustrated in figure 9 by plotting $t_{\hat{\sigma}}$, $\hat{\sigma}$, and t_{CH} (figure 9 (a), (b), and (c) respectively) against the normalised coordinates r_f , θ_f , and z_f (given in equations (18) to (20)). These plots demonstrate the wide range thermal stress responses recorded and verify that, while there are some general trends visible, critical parameter values are dependent on header geometry and analysis point coordinates.

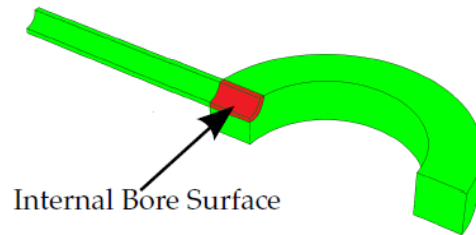


Figure 7: Illustration of the sample region in the vicinity of the stub penetration (highlighted in red).

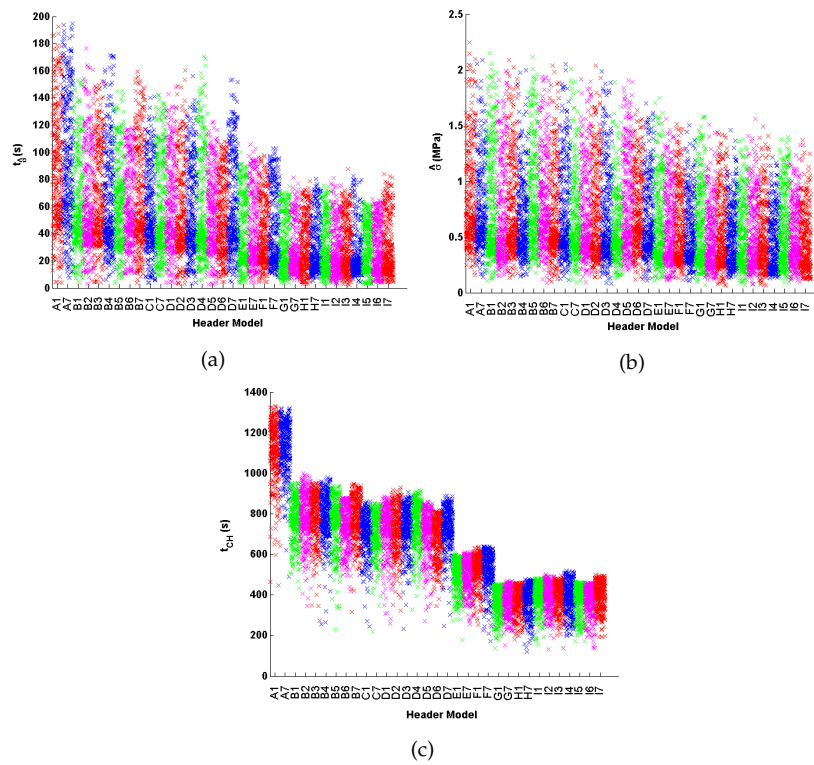


Figure 8: Critical parameters determined from FEA results (note alternating colours are used to distinguish different header geometries only), showing (a) $t_{\hat{\sigma}}$, (b) $\hat{\sigma}$, and (c) t_{CH} .

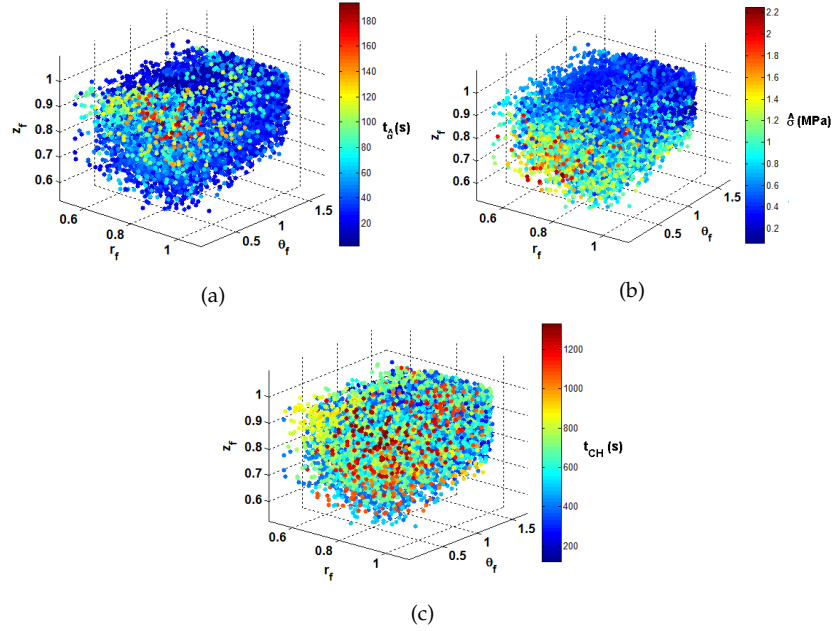


Figure 9: Variation of critical parameters ((a) $t_{\hat{\sigma}}$, (b) $\hat{\sigma}$, and (c) t_{CH}) with normalised coordinates r_f , θ_f , and z_f (determined from FEA results).

4.2. Neural network results

The process of training a neural network is an attempt to optimise the weight and bias values associated with each neuron in the network. The inherent danger in using approximate functions to describe a behaviour is that, given enough complexity, any response can be estimated without actual trends being recognised. In order to guard against this “over fitting”, validation cases are often considered. The performance of the network is therefore judged not only against results used to train the network, but also against results that the network has never “seen”. Header models used for training and validation were determined from the total set randomly (the models used for each purpose may be determined from figures 10 to 12, noting validation models are highlighted).

The predictive capability of the developed network is generally very good. Figures 10 to 12 illustrate this for the $t_{\hat{\sigma}}$, $\hat{\sigma}$, and t_{CH} critical parameters, respectively. Sub-figures (a) and (c) compare the absolute critical parameter values determined from FEA (true) and the neural network method (predicted, designated neural network) for the training and validation processes, respectively. In the interest of clarity, sparse and randomly sampled versions of sub-figures (a) and (c) are presented in sub-figures (b) and (d), respectively. Entire data sets are reported in sub-figures (a) and (c) in order to illustrate general trends in the data, with discernible example observed/predicted data pairs presented in sub-figures (b) and (d). Sub-figures (e) and (f) demonstrate the frequency of

errors recorded between FEA (true) values for the critical parameters and the associated neural network predictions. For training cases, at least 70% of the results of have an error of 5% or less, with approximately 90% of cases have errors of 10% or less. Comparatively for the validation cases, at least 55% of the results of have an error of 5% or less, with approximately 80% of cases have errors of 10% or less.

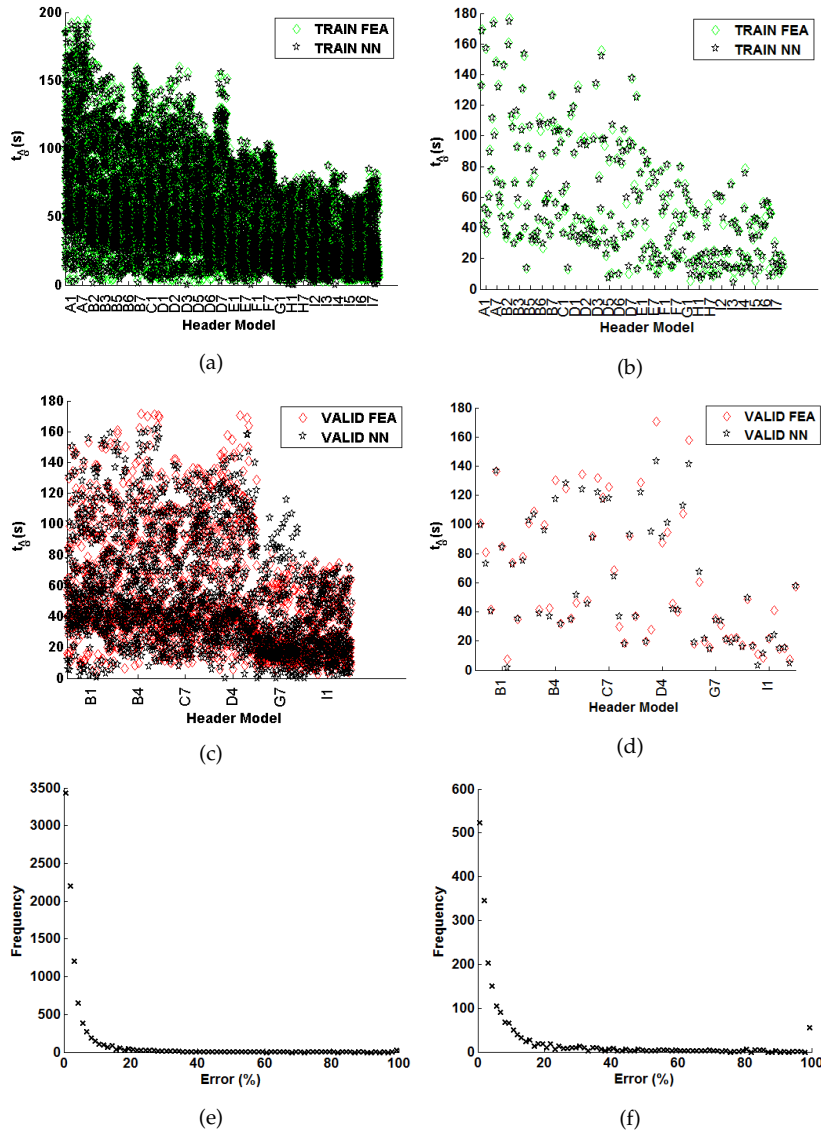


Figure 10: Neural network performance for prediction of t_{δ} , showing (a) the prediction of absolute values for the training cases, (b) a random sample of observed and predicted values for the training set, (c) the prediction of absolute values for the validation cases, (d) a random sample of observed and predicted values for the validation set, (e) error frequency distributions for the training cases, and (f) error frequency distributions for the validation cases.

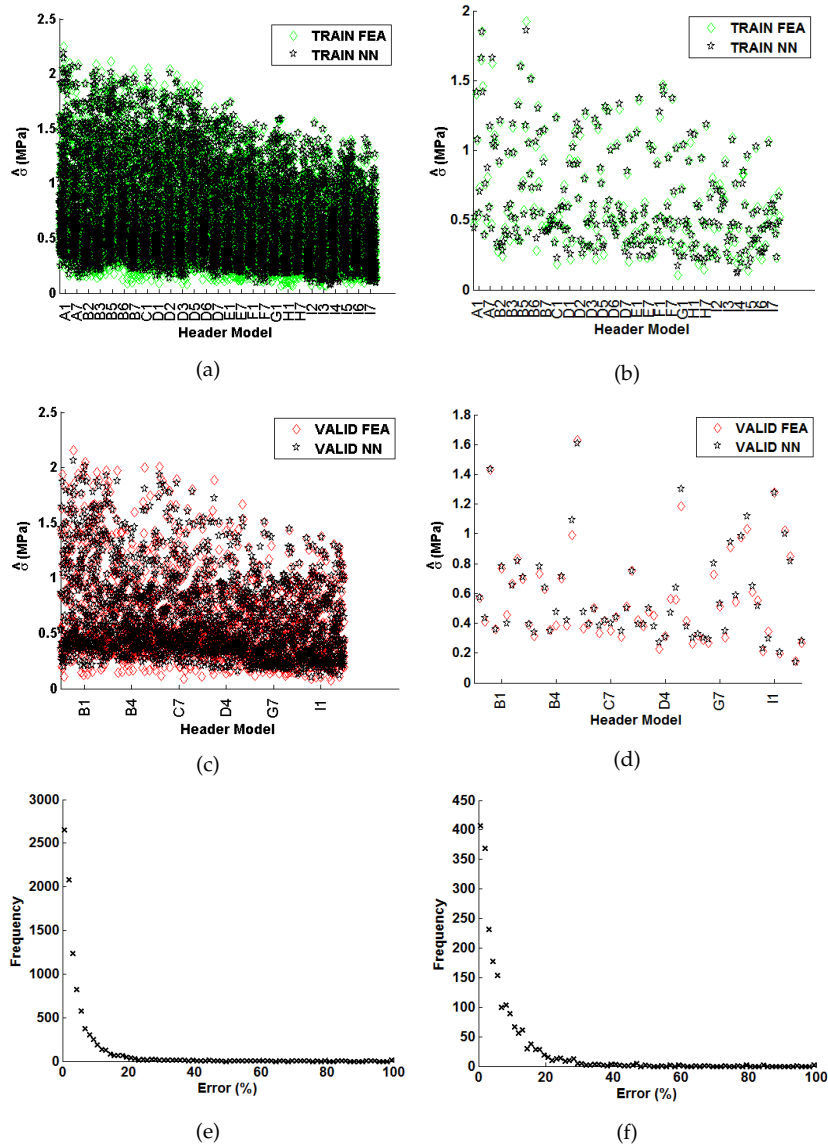


Figure 11: Neural network performance for prediction of $\hat{\sigma}$, showing (a) the prediction of absolute values for the training cases, (b) a random sample of observed and predicted values for the training set, (c) the prediction of absolute values for the validation cases, (d) a random sample of observed and predicted values for the validation set, (e) error frequency distributions for the training cases, and (f) error frequency distributions for the validation cases.

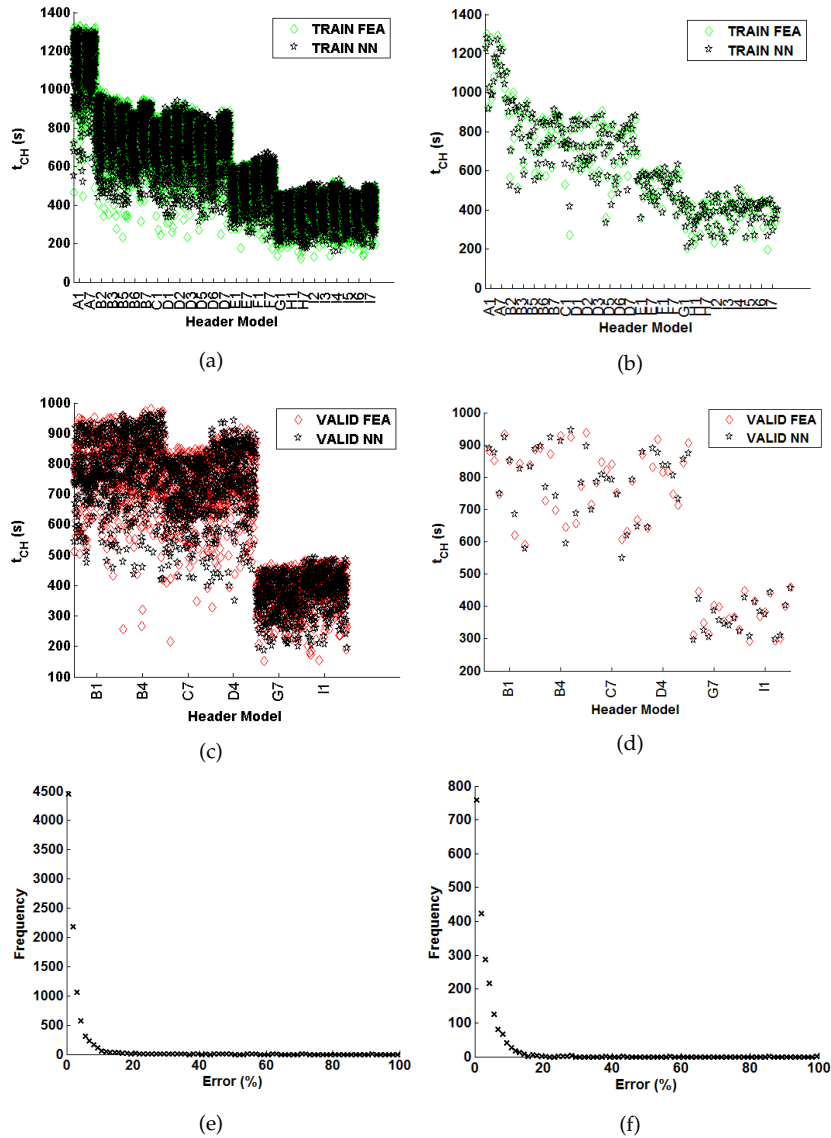


Figure 12: Neural network performance for prediction of t_{CH} , showing (a) the prediction of absolute values for the training cases, (b) a random sample of observed and predicted values for the training set, (c) the prediction of absolute values for the validation cases, (d) a random sample of observed and predicted values for the validation set, (e) error frequency distributions for the training cases, and (f) error frequency distributions for the validation cases.

4.3. Stress profile reconstruction and example applications

To demonstrate the application of the developed network beyond the cases used for training and validation, ten additional header models were created (as shown in table 4, hereafter referred to as “test” cases). Reconstructed stress profiles for these test cases are considered here (examples may be seen in figure 13). It was mentioned previously that broad relationships between critical parameter values may be observed (see section 3.2). These relationships provide a concise summary of the critical parameters and examples are provided for cases T4, T9, and T10 in figures 14 to 16, respectively.

Cracks initiating at the internal bore surface (see figure 7) of the stub penetration is a significant concern for thick walled components such as headers. Thermal stresses tend to be highest at the inner shell location of this surface (for monotonic thermal loads). In order to demonstrate the predictive capability of the developed network for a typical application, thermal stress contour plots are compared on this surface for three example headers from the test set (see figure 17). Results are taken from unit temperature step analyses using both the “true” FEA and “approximate” neural network approaches. Good predictions for these cases suggest that good estimations will be achieved once Green’s functions are integrated for arbitrary thermal loads. Time instances are chosen such that peak thermal stresses are the highest observed in the analyses.

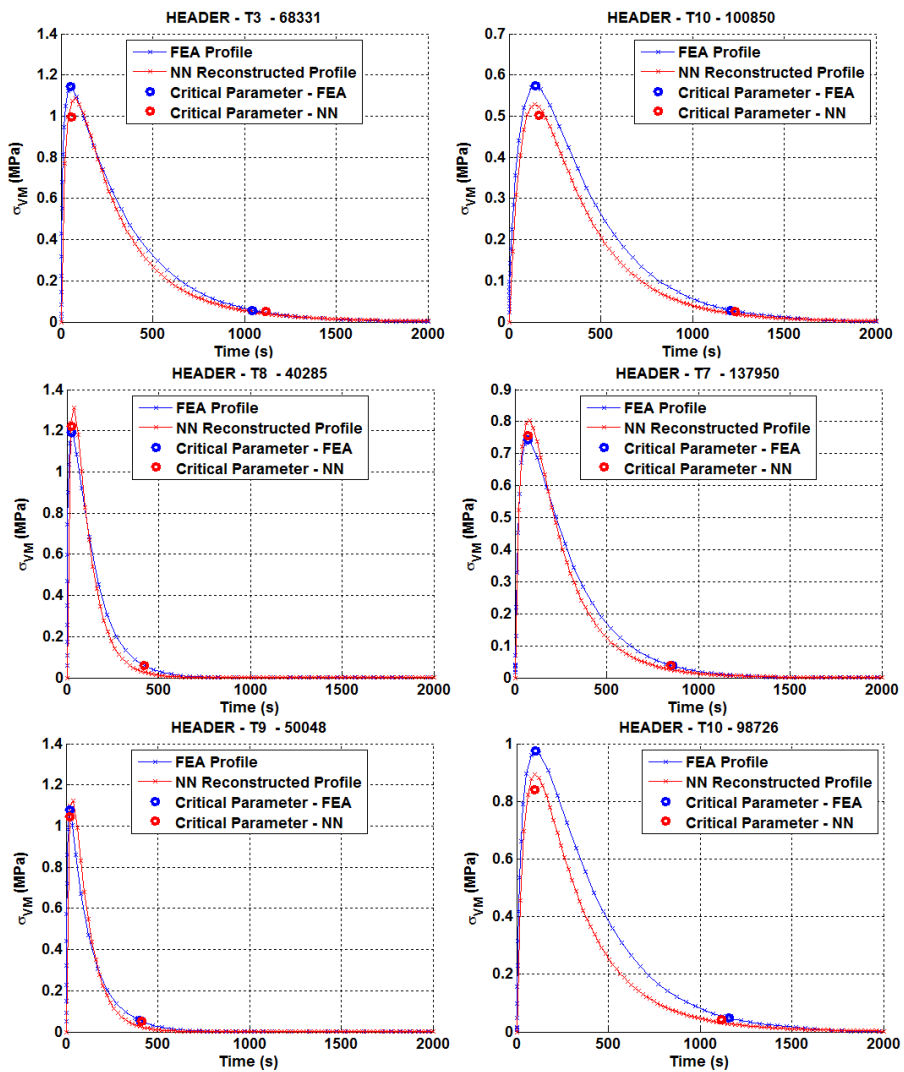


Figure 13: Example reconstructed stress profiles.

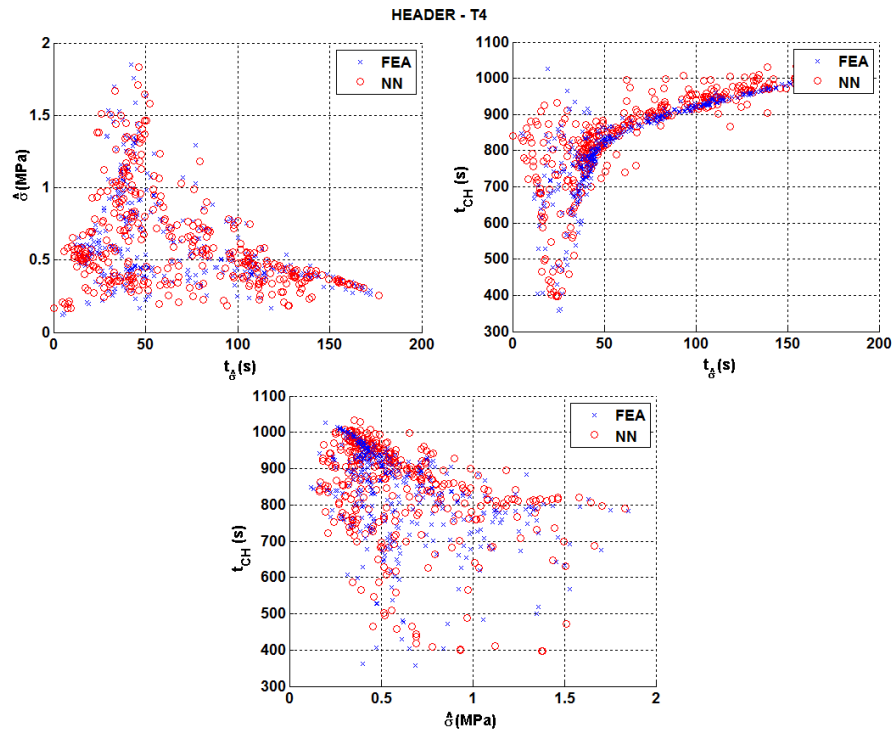


Figure 14: The relationships between critical parameters for test case T4, showing true (FEA) and predicted (neural network - NN) results.

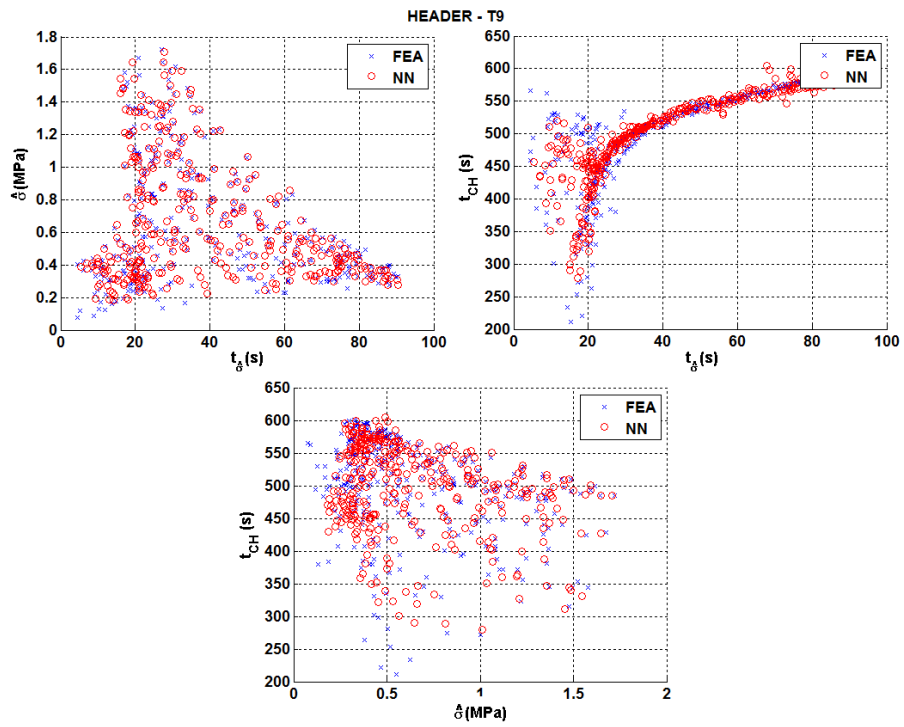


Figure 15: The relationships between critical parameters for test case T9, showing true (FEA) and predicted (neural network - NN) results.

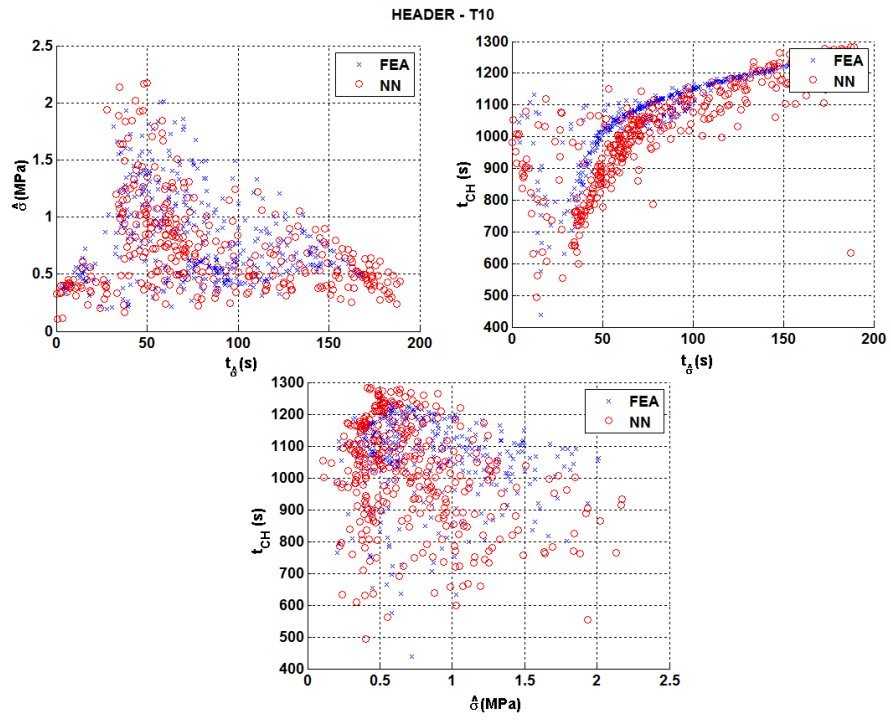


Figure 16: The relationships between critical parameters for test case T10, showing true (FEA) and predicted (neural network - NN) results.

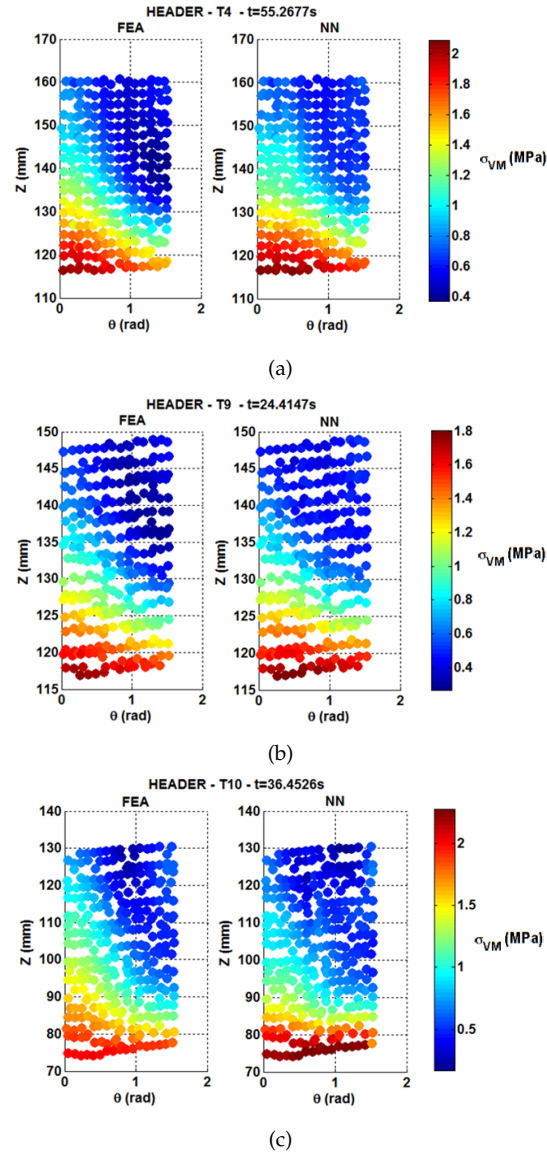


Figure 17: Predicted stress contour plots (using the “true” FEA and “approximate” neural network approaches) on the internal bore surfaces for header models (a) T4, (b) T9, and (c) T10.

5. Discussion

The performance of the neural network approach will be quantified in the following section. Several parameters are determined for the training/validation

and test model sets in figures 18 and 19, respectively. Calculated parameters are mean coefficients of determination R^2 (averaged over all reconstructed unit temperature step stress profile, shown in sub-figure (a))[40], mean peak stress difference (sub-figure (b)), mean error (sub-figure (c)), and the standard deviation of R^2 (sub-figure (d)). Mean error values are determined by comparing integrals of the stress functions for the true (FEA) and predicted (neural network) methods.

Generally, training, validation and test results return similar R^2 , stress difference and error values, suggesting that over fitting has been avoided in the network. R^2 values are typically high (over 0.92) with small standard deviations (≈ 0.1). Mean peak stress differences for the reconstructed stress profiles are also commonly small at $0.14MPa$ compared to maximum thermal stresses which are usually greater than $2MPa$. Mean errors confirm this observation, being typically less than 10%. Comparatively large errors ($\approx 18\%$) and low R^2 (≈ 0.8) are noted in some cases, namely H1 (training set), G7 (validation set), and T3 and T4 (test set). Figure 20 illustrates the distribution of the modelled header geometries using the ratios ϕ_{STUB}/t_{STUB} and ϕ_{SHELL}/t_{SHELL} and highlights a particular model's function in the network development (training, validation or testing). It can be seen from this plot that the four models mentioned previously lie on the edges of the simulated range. Future work should therefore look to extend results for $\phi_{SHELL}/t_{SHELL} > 14$. For a large section of the industrially relevant dimensions however, the developed neural network can be demonstrated to provide excellent estimates of unit thermal stress profiles.

The accuracy of the neural network method may be further investigated using the results given in table 5, wherein mean squared errors (MSE , see equation (23)) are determined for the training, validation, test data sets. It is noted here that magnitudes of MSE are compatible to the average critical parameter values for $t_{\hat{\sigma}}$ and t_{CH} . Given that good levels of agreement were suggested by the alternative metrics, this result suggest that several significant errors exist in the predictive data set. Future work will look to minimise these, however the general predictive ability of the neural network method reported in the results section should still be noted. It is possible that future implementations of this method will implement constraints (such as continuity) on the predicted stress field in order to identify outliers. A simple non-linear regression model is also presented here in order to establish what level of predictive capability can be achieved with a minimal amount of effort. The regression model takes the form of equation (24), where y is a predicted parameter ($t_{\hat{\sigma}}$, $\hat{\sigma}$, or t_{CH}), x_i is one of seven input parameters (ϕ_{SHELL} , t_{SHELL} , ϕ_{STUB} , t_{STUB} , r_f , θ_f , and z_f), and b_k is a set of 15 fitting parameters (determined against the training data set using the MATLAB function *fitnlm*, resulting R^2 values are reported for comparison). MSE values are orders of magnitude greater for all critical parameters in table 5. While correlations of course exist between the input and output parameters, finding an expression of similar form to equation (24) is likely to be a laborious task. The advantage the neural network method may be further examined by considering the level of correlation between the input parameters and the outputs. Spearman's rank correlation coefficient (see equation (25), where r_s is

the correlation coefficient, d is a difference between parameter pairs, and n is the number of parameter pairs) is applicable for non-normal distributions and can indicate the existence of non-linear correlations. Correlation parameters and p-values were determined for the entire population of FEA results using the MATLAB function *corr* and are given in table 6. Almost all correlations are statistically significant (p-values are typically less than a standard 0.05 significance level), however many of the correlations are evaluated as being weak (i.e. a small r_s values are observed).

$$y \approx b_0 + \sum_{i=1}^7 b_{2i-1} x_i^{b_{2i}} \quad (24)$$

$$r_s = 1 - \frac{6\sum d^2}{n^3 - n} \quad (25)$$

Table 5: Mean square error (MSE) evaluations for the developed neural network method and a simple non-linear regression model.

Data Set		NN-MSE	N-L Regression - MSE	N-L Regression - R^2
Training	$t_{\hat{\sigma}}$	2.88	891.03	0.23
	$\hat{\sigma}$	7.52×10^{-4}	4.70×10^{-2}	0.62
	t_{CH}	415.50	5.81×10^3	0.89
Validation	$t_{\hat{\sigma}}$	122.65	1.14×10^3	-
	$\hat{\sigma}$	1.80×10^{-3}	5.01×10^{-2}	-
	t_{CH}	767.03	5.42×10^3	-
Test	$t_{\hat{\sigma}}$	322.23	1.67×10^3	-
	$\hat{\sigma}$	1.01×10^{-2}	6.65×10^{-2}	-
	t_{CH}	3.03×10^3	9.48×10^3	-

Table 6: Spearman’s rank correlation coefficient evaluations and associated p-values, showing the significance of correlations found between the “input” parameters (from the neural network’s perspective, namely ϕ_{SHELL} , t_{SHELL} , ϕ_{STUB} , t_{STUB} , r_f , θ_f , and z_f) and the “output” parameters ($t_{\hat{\sigma}}$, $\hat{\sigma}$, and t_{CH}).

	$t_{\hat{\sigma}}$	$\hat{\sigma}$	t_{CH}
	r_s /p-value	r_s /p-value	r_s /p-value
ϕ_{SHELL}	$0.16/5.38 \times 10^{-88}$	$0.13/1.38 \times 10^{-58}$	$0.24/2.75 \times 10^{-194}$
t_{SHELL}	0.50/0.00	$0.26/2.43 \times 10^{-234}$	0.88/0.00
ϕ_{STUB}	$-0.07/1.72 \times 10^{-18}$	$-0.03/1.45 \times 10^{-4}$	$-0.13/2.81 \times 10^{-56}$
t_{STUB}	$-0.02/4.26 \times 10^{-2}$	$-0.11/3.70 \times 10^{-45}$	$0.01/7.15 \times 10^{-1}$
r_f	$0.22/2.99 \times 10^{-166}$	$-0.21/5.49 \times 10^{-150}$	$0.08/8.42 \times 10^{-22}$
θ_f	-0.49/0.00	$-0.28/5.78 \times 10^{-270}$	$-0.24/2.82 \times 10^{-199}$
z_f	$0.15/1.14 \times 10^{-78}$	-0.54/0.00	$0.03/8.15 \times 10^{-4}$

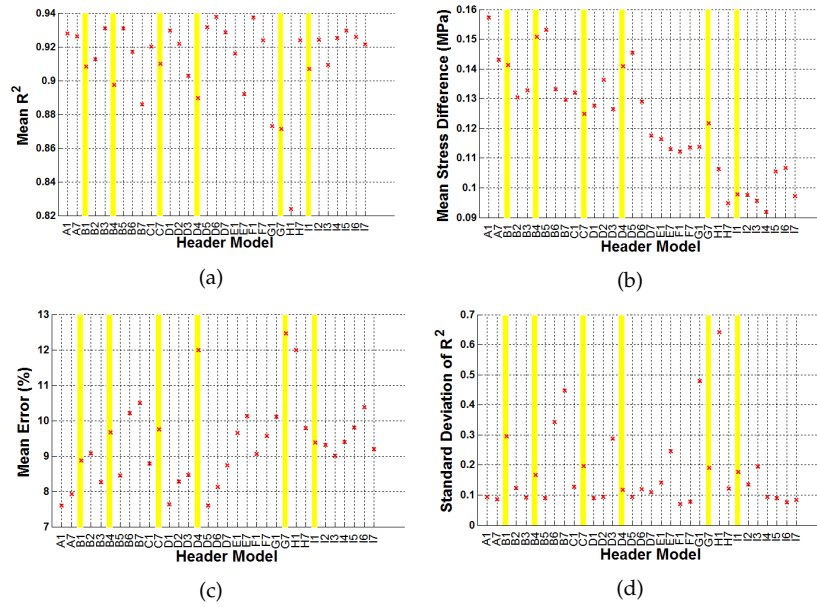


Figure 18: Comparative performance of the developed neural network for the training and validation header sets (validation cases are highlighted), showing (a) mean coefficients of determination, (b) mean peak stress differences, (c) mean error, and (d) the standard deviation of the coefficient of determination.

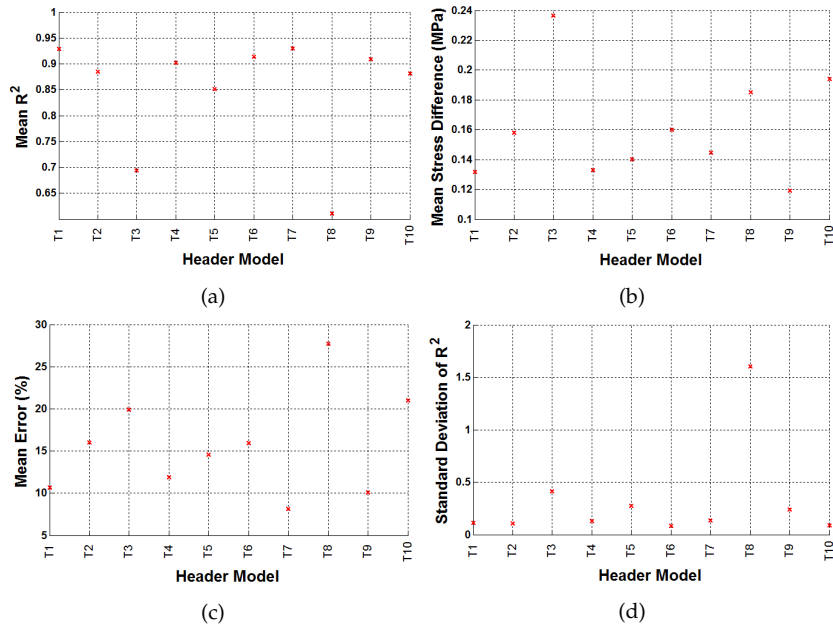


Figure 19: Comparative performance of the developed neural network for the test header set, showing (a) mean coefficients of determination, (b) mean peak stress differences, (c) mean error, and (d) the standard deviation of the coefficient of determination.

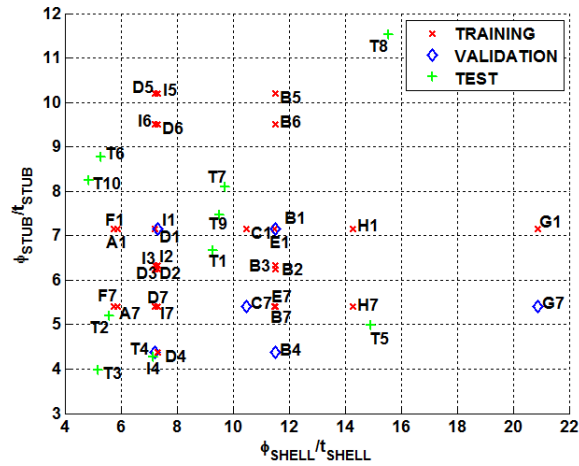


Figure 20: The distribution of modelled header geometries, expressed through the ratios ϕ_{STUB}/t_{STUB} and ϕ_{SHELL}/t_{SHELL} .

A good level of agreement is observed between stress contour plots gener-

ated through uncoupled thermal stress FEA analyses and using the approximate neural network approach for a wide range of header geometries (see figure 17 for example predictive contour plots). The predictive capability of the network developed here may be demonstrated by considering statistical measures of the training/validation/test data sets and observing how these differ for the FEA (“true”) and neural network (predicted) results. Histogram distributions are presented for the critical parameters $t_{\hat{\sigma}}$, $\hat{\sigma}$, and t_{CH} in figures 21 to 23, respectively. In each case the entire population (all available data), training, validation, and test data sets are considered in order to establish any similarities between the data sets. Weibull probability density functions (PDFs), expressible for positive x values by equation (26), have also been fitted to each distribution. PDF parameters, namely the scale factor λ and the shape factor k , are reported in tables 7 to 9 for $t_{\hat{\sigma}}$, $\hat{\sigma}$, and t_{CH} , respectively. While the Weibull distribution may not be appropriate for the description of t_{CH} (see figure 23) as the distribution is clearly not uni-modal, it is applied nether the less so that comparisons between the four data sets may be made. In all cases, distributions observed in the FEA data sets are usually accurately reproduced in the neural network equivalents, further suggesting that the large MSE values reported in table 5 are due to outliers rather than some systematic error in the neural network analysis method. Of particular interest is the differences between the test data set distributions and the training/validation sets. Such observations support the suggestion made previously that larger errors in the test data set are due to the consideration of a different region of parameter space. Future work should look to perform such distribution comparisons prior to assigning training/validation/test identities.

$$f(x; \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k} \quad (26)$$

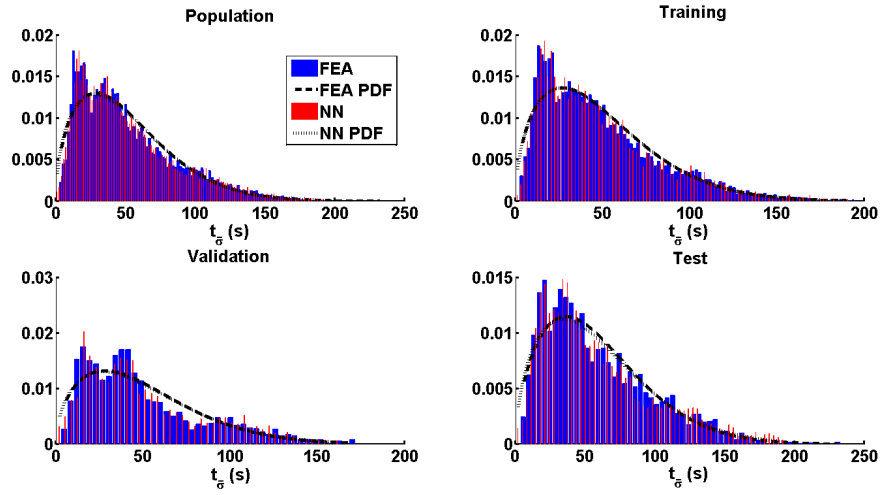


Figure 21: Probability density functions calculated for distributions of $t_{\hat{\sigma}}$, showing results for the full population and the training, validation, and test subsets. FEA (“true”) and neural network predicted results are presented.

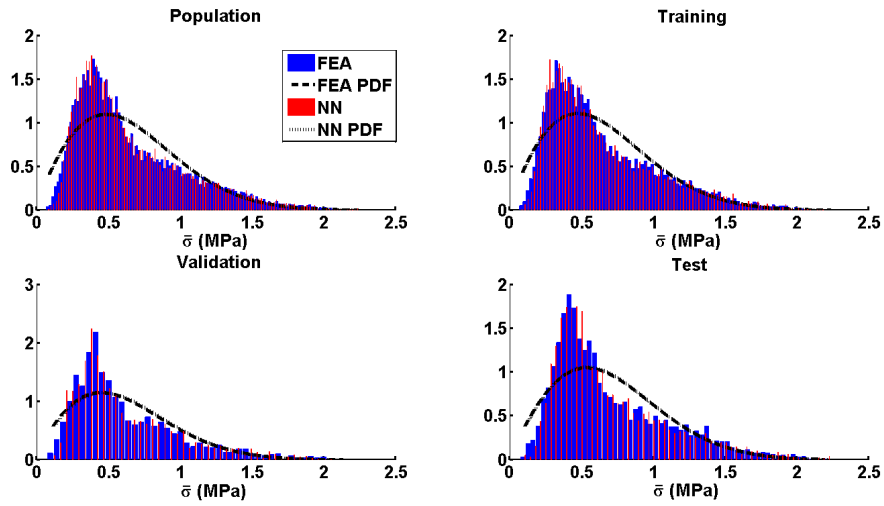


Figure 22: Probability density functions calculated for distributions of $\hat{\sigma}$, showing results for the full population and the training, validation, and test subsets. FEA (“true”) and neural network predicted results are presented.

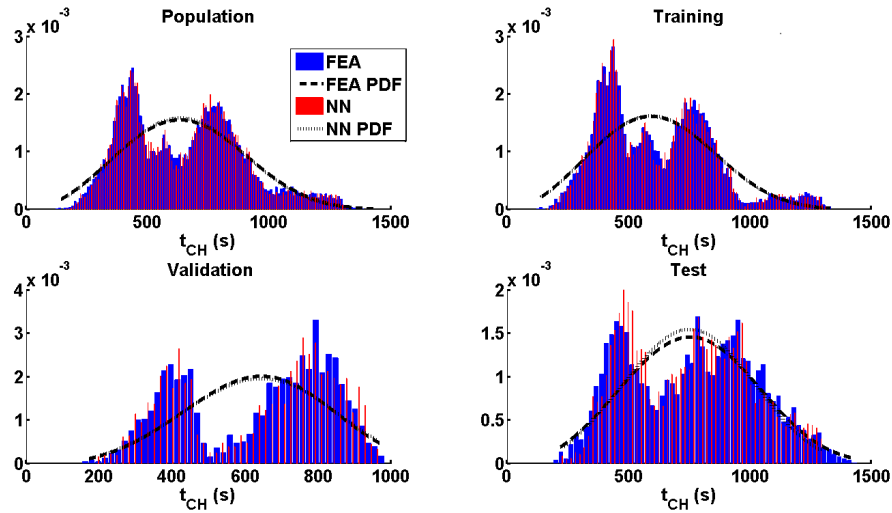


Figure 23: Probability density functions calculated for distributions of t_{CH} , showing results for the full population and the training, validation, and test subsets. FEA (“true”) and neural network predicted results are presented.

Table 7: Weibull distribution parameters calculated for the t_{δ} data sets, see figure 21.

Data Set	FEA		NN	
	λ	k	λ	k
Population	57.65	1.52	57.76	1.50
Training	54.76	1.51	54.65	1.51
Validation	56.49	1.52	56.98	1.49
Test	66.35	1.60	66.88	1.54

Table 8: Weibull distribution parameters calculated for the δ data sets, see figure 22.

Data Set	FEA		NN	
	λ	k	λ	k
Population	0.74	1.83	0.74	1.85
Training	0.73	1.82	0.73	1.83
Validation	0.70	1.81	0.71	1.84
Test	0.79	1.90	0.79	1.92

Table 9: Weibull distribution parameters calculated for the t_{CH} data sets, see figure 23.

Data Set	FEA		NN	
	λ	k	λ	k
Population	731.12	2.88	728.49	2.94
Training	691.61	2.82	691.45	2.84
Validation	701.29	3.69	704.84	3.60
Test	848.98	3.18	836.81	3.34

The developed technique is of interest to industry and will be used to inform the future operation of simulated components. Advancements may be made in several areas however. Typically, fatigue analysis will depend upon study of the complete stress tensor. Only von Mises stresses are presented in the present work by way of example, however other stress components (steam chests, for example) typically have similar thermal stress profiles to the ones illustrated, suggesting that the method may be easily extended to predict all stress components. In order to aid in reader comprehension, an example itemised implementation of the analysis method alluded to here is presented:

- Historic data (steam temperatures and pressures) is extracted from power plant records for the analysis component.
- Historic data is sampled to create a discrete time series.
- Component geometry data (pressure vessel diameters and wall thicknesses, for example) are used as inputs to generate, from a trained neural network, a set of Green's functions for that component. Separate Green's functions are generated for different spacial positions in the component (the frequency of which depending on the level of detail required in the analyses).
- For each spatial position in the analysis component, unit thermal stress responses (given by the Green's functions) are scaled for each temperature increment in the discrete time series and summed, giving an approximation of the transient thermal stress response at that spatial position.
- Transient thermal stress responses are superimposed on stress fields related to internal pressure (which may be easily approximated) and corrected using, say, Neuber's rule.
- Cycle counting methods are implemented over the time series in order to determine the total number of closed loading cycles in the transient stress history and remnant fatigue life is estimated.

To limit the complexity of the analysis, a simple single stud penetration header model has been assumed. This is not truly representative of real world

components which will typically contain several penetrations on a particular “row”, however it does provide useful first estimates of the stress field (which may be all that is required). The effects of multiple stub penetrations and the angles between these penetrations will be investigated in the future. Several developments could also be made in the stress profile reconstruction method. At present, single values are chosen for the parameters f_1 and f_2 (used to estimate intermediate stress values based on critical parameter values). These could however be functions of the estimated critical parameters, potentially providing more representative intermediate points prior to smoothing. Alternatively, conditions for smooth and continuous stress/temperature fields may be translated into the predicted surfaces through appropriate constraints. In this way, perturbations due to errors in the approximate neural network method may be minimised. The general methodology applied in the present work may also be extended to other components where geometry may be simplified/generalised.

6. Conclusions

Determining transient thermal stress magnitudes is important for the safe operation of thick walled components due to the potentially life limiting effects of thermal fatigue. Estimating these stresses can be expensive however due to the wide variety of thermal and mechanical loads that may be applied to a component. While mechanical loads can be accounted for with relative ease for linear elastic cases by scaling results from some reference case, thermal loads (as a result of their transient nature) typically require a time dependent temperature field to be found which is then used to estimate thermal expansion and the generation of internal forces (in uncoupled simulations). The Green’s function method allows reference unit temperature step stress solutions to be integrated over time at a particular analysis point in a structure to estimate thermal stress distributions. This removes the need for bespoke analyses for a particular structure geometry, however Green’s functions will be necessarily dependent upon the structure’s geometry and the position of the analysis point. The present work generalises the effects of geometry and analysis point position through the use of a neural network. This allows reference Green’s functions to be generated from geometry data alone (assuming a reference material) for thick walled header components that are typically employed in the power industry. Multiple component geometries have been tested in the present work, including several that were not originally considered in the development/training of the network (thus providing protection against over fitting). Transient stress profiles typically have coefficients of determination greater than 0.92 and error less than 10%. Future work should look to extend the range of geometries suitable for analysis using the neural network method by increasing the data available for $\phi_{SHELL}/t_{SHELL} > 14$. Of particular note in the present work is the identification three “critical parameters” which may be used to construct an approximation of the thermal stress Green’s function. This greatly simplifies the parametrisation of the Green’s function and enables techniques such as neural network approximation to be employed effectively.

7. Acknowledgements

The authors would like to sincerely thank EDF UK for their willingness to share information and for funding the presented research.

References

- [1] Samal MK, Dutta BK, Guin S, Kushwaha HS. A Finite Element Program for on-Line Life Assessment of Critical Plant Components. *Engineering Failure Analysis*. 2009;16:85–111.
- [2] Mukhopadhyay NK, Dutta BK, Kushwaha HS. On-Line Fatigue-Creep Monitoring System for High-Temperature Components of Power Plants. *International Journal of Fatigue*. 2001;23:549–560.
- [3] Daga R, Samal MK. Real-Time Monitoring of High Temperature Components. In: 6th International Conference on Creep, Fatigue and Creep-Fatigue Interaction. vol. 63. *Metallurgical Journal*; 2012. p. 87–89.
- [4] Benatoa A, Braccoc S, Stoppatoa A, Mirandola A. LTE: A Procedure to Predict Power Plants Dynamic Behaviour and Components Lifetime Reduction During Transient Operation. *Applied Energy*. 2016;162:880–891.
- [5] Wermelinger T, Bruckmüller F, Heinz B. Fatigue Monitoring in the Context of Long-Term Operation of the Goesgen Nuclear Power Plant Using AREVA's FAMOSi. In: ASME 2015 Pressure Vessels and Piping Conference Proceedings, Vol. 7: Operations, Applications and Components. doi:10.1115/PVP2015-45272.; 2015. .
- [6] Heinz B, Wu D. AREVA's Modularized Fatigue Monitoring for Lifetime Extension and Flexible Plant Operation. In: ASME 2014 Pressure Vessels and Piping Conference Proceedings, Vol. 7: Operations, Applications and Components. doi:10.1115/PVP2014-28726.; 2014. .
- [7] Farraghera TP, Scully S, O'Dowd NP, Leen SB. Development of life assessment procedures for power plant headers operated under flexible loading scenarios. *International Journal of Fatigue*. 2013;49:50–61.
- [8] Paterson IR, Wilson JD. Use of Damage Monitoring Systems for Component Life Optimisation in Power Plant. *International Journal of Pressure Vessels and Piping*. 2002;79:541–547.
- [9] Ainsworth RA. R5 Procedures for Assessing Structural Integrity of Components Under Creep and Creep-Fatigue Conditions. *International Materials Reviews*. 2006;51:107–126.
- [10] Ainsworth RA, Booth SE. Use of R5 in Plant Defect Assessment. *Materials at High Temperatures*. 1998;15:299–302.
- [11] Ainsworth RA, Hooton DG. R6 and R5 Procedures: The Way Forward. *International Journal of Pressure Vessels and Piping*. 2008;85:175–182.
- [12] Sakaia K, Hojob K, Katoc A, Umeharac R. On-Line Fatigue-Monitoring System for Nuclear Power Plant. *Nuclear Engineering and Design*. 1994;153:19–25.

- [13] Maekawaa O, Kanazawaa Y, Takahashib Y, Tanic M. Operating Data Monitoring and Fatigue Evaluation Systems and Findings for Boiling Water Reactors in Japan. *Nuclear Engineering and Design*. 1995;153:135–143.
- [14] Stevens GL, Ranganath S. Use of On-Line Fatigue Monitoring of Nuclear Reactor Components as a Tool for Plant Life Extension. *Journal of Pressure Vessel Technology*. 1991;113:349–357.
- [15] Boley BA, Weiner JH. *Theory of Thermal Stresses*. John Wiley and Sons, Inc.; 1960.
- [16] Cowan JD. Discussion: McCulloch-Pitts and Related Neural Networks from 1943 to 1989. *Bulletin of Mathematical Biology*. 1990;52:73–97.
- [17] Yun GJ, Ghaboussi J, Elnashai AS. A New Neural Network-Based Model for Hysteretic Behavior of Materials. *International Journal for Numerical Methods in Engineering*. 2008;73(4):447–469.
- [18] Man H, Furukawa T. Neural Network Constitutive Modelling for Non-Linear Characterization of Anisotropic Materials. *International Journal for Numerical Methods in Engineering*. 2011;85(8):939–957.
- [19] Venkatesh V, Rack HJ. Neural Network Approach to Elevated Temperature Creep-Fatigue Life Prediction. *International Journal of Fatigue*. 1999;21(3):225–234.
- [20] Hambli R. Apparent Damage Accumulation in Cancellous Bone Using Neural Networks. *Journal of the Mechanical Behavior of Biomedical Materials*. 2011;4(6):868–878.
- [21] Jin H, Wu S, Peng Y. Prediction of Contact Fatigue Life of Alloy Cast Steel Rolls Using Back-Propagation Neural Network. *Journal of Materials Engineering and Performance*. 2013;22(12):3631–3638.
- [22] Kao CS, Yeh IC. Using Neural Networks to Integrate Structural Analysis Package and Optimization Package. *Neural Computing and Applications*. 2016;27:571–583.
- [23] Quéau LM, Kimiaei M, Randolph MF. Artificial Neural Network Development for Stress Analysis of Steel Catenary Risers: Sensitivity Study and Approximation of Static Stress Range. *Applied Ocean Research*. 2014;48:148–161.
- [24] Adeli H, Jiang X. *Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures*. 1st ed. Boca Raton, Florida, U.S.A.: CRC Press; 2009.
- [25] Alavala CR. *Fuzzy Logic and Neural Networks: Basic Concepts and Application*. New Age International; 2008.

- [26] Chapra SC, Canale RP. Numerical Methods for Engineers. 6th ed. Boston, Massachusetts, U.S.A.: McGraw-Hill Higher Education; 2010.
- [27] Canning A, Gardner E. Partially Connected Models of Neural Networks. *Journal of Physics A: Mathematical and General*. 1988;21:3275–3284.
- [28] Weymaere N, Martens JP. On the Initialization and Optimization of Multilayer Perceptrons. *IEEE Transactions on Neural Networks*. 1994;5:738 – 751.
- [29] Duch W, Adamczak R, Jankowski N. Initialization and Optimization of Multilayered Perceptrons. In: *Proceedings of the 3rd Conference on Neural Networks and their Applications*. Kule, Poland; 1997. .
- [30] Jamett M, Acuna G. An Interval Approach for Weight’s Initialization of Feedforward Neural Networks. In: Gelbukh A, Reyes-Garcia C, editors. *MICAI 2006: Advances in Artificial Intelligence*. vol. 4293 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg; 2006. p. 305–315.
- [31] Vitela J, Reifman J. Premature Saturation in Backpropagation Networks: Mechanism and Necessary Conditions. *Neural Networks*. 1997;10:721 – 735.
- [32] Li G, Alnuweiri H, Wu Y. Acceleration of Back Propagations Through Initial Weight Pre-Training with Delta Rule. In: *IEEE International Conference on Neural Networks*. San Francisco, U.S.A.; 1993. p. 580 – 585.
- [33] Nguyen D, Widrow BI. Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights. In: *Proceedings of the International Joint Conference on Neural Networks*. vol. 3. Washington, D. C., U.S.A.; 1990. p. 21–26.
- [34] Yoon HS, Bae CS, Min BW. Neural Networks using Modified Initial Connection Strengths by the Importance of Feature Elements. In: *IEEE International Conference on Systems, Man and Cybernetics: Intelligent Systems for the 21st Century*. Vancouver, Canada; 1995. p. 458 – 461.
- [35] *Optimisation Toolbox TM 4 User’s Guide*. T. M. Inc; 2008.
- [36] *MATLAB Neural Network Toolbox User’s Guide (R2014a)*. Mathworks; 2014.
- [37] *ABAQUS 6.9 Online Documentation*. Dassault Systeme’s; 2009.
- [38] Yaghi A, Hyde TH, Becker AA, Williams JA, Sun W. Residual Stress Simulation in P91 Pipe Welds. *Journal of Materials Processing Technology*. 2005;167:480–487.
- [39] Nakoneczny GJ. Life Assessment of High Temperature Headers. In: *57th American power Conference*. vol. 63. *Metallurgical Journal*; 1995. p. 87–89.
- [40] *Navidi W. Statistics for Engineers and Scientists*. McGraw-Hill; 2008.